The role of knowledge variety and intensity for regional innovation

Sam Tavassoli · Nunzia Carbonara

Abstract This paper analyses the effect of variety and intensity of knowledge on the innovation of regions. Employing data for Swedish functional regions, the paper tests the role of the variety (related and unrelated) and intensity of (1) internal knowledge generated within the region and also (2) external knowledge networks flowing into the region in explaining regional innovation, as measured by patent applications. The empirical analysis provides robust evidence that both the variety and intensity of internal and external knowledge matter for regions' innovation. When it comes to variety, related variety of knowledge plays a superior role.

Keywords Knowledge intensity · Knowledge variety · Related variety · Unrelated variety · Internal knowledge · External knowledge · Patent applications · Functional regions

JEL Classifications O32 · F14 · R12 · L26

S. Tavassoli (✉)
Department of Industrial Economics, Blekinge Institute of Technology (BTH), 371 79 Karlskrona, Sweden
e-mail: sam.tavassoli@bth.se

N. Carbonara
Department of Mechanical and Management Engineering, Politecnico di Bari, Viale Japigia, Bari, Italy
e-mail: ncarbonara@poliba.it

1 Introduction

The role of knowledge[1] is considered as crucial for the innovation of regions. But is it the 'internal knowledge' generated within the region that matter or is it 'external knowledge' brought into the region through trade networks? Moreover, considering either of these two intra/extra-regional sources of knowledge, is it the 'intensity' of the knowledge that matters or is it rather the 'variety' of knowledge? These questions have recently received attention in the literature, but few studies try to address them in a common empirical setting. Some are still unexplored, such as the role of 'variety' (related and unrelated variety) for regional innovativeness, and are still open both in the traditional and in the more recent literature (see Beaudry

---

[1] In this paper, "knowledge" refers to all three types of knowledge classified by Karlsson and Johansson (2006): scientific knowledge, technological knowledge, and entrepreneurial knowledge. Scientific knowledge has the character of a pure public good, although it is generally only available to those with the relevant scientific training. Technological and entrepreneurial knowledge are non-rivalrous and partially excludable goods, where the latter is often the result of learning-by-doing. All three types are argued to be patentable and there is indeed evidence on increasing propensity to patent all three types. Since the phenomenon under investigation in this paper is actually patent application, it seems plausible to include all of them as the conceptualization of knowledge.

and Schiffauerova 2009; De Groot et al. 2009).[2] This paper directly bears on these questions, and provides an empirical analysis of the role of variety and intensity of internal and external knowledge for regional innovation.

A common argument in the literature is indeed that innovation depends on the 'intensity' and 'variety' of knowledge (Audretsch and Vivarelli 1996; Audretsch and Feldman 2004; Beaudry and Schiffauerova 2009), reflecting a general shift from cost- to knowledge-based models of regional growth. Knowledge-based theories of regional growth and innovation (see for example Maskell 2001; Maskell and Malmberg 2002) emphasize the nature of local knowledge (Tallman et al. 2004), the intensity (and frequency) of knowledge transfer processes among local firms (Gordon and Mccann 2000; Mesquita 2007), and the variety of knowledge in the region (Jacobs 1969; Glaeser et al. 1992; Frenken et al. 2007). It is further argued that the process of knowledge generation is not only activated by means of interactions between firms located *inside* regions. It can also be brought into the region *from outside*, through international trade networks (Castellani 2002; Bathelt et al. 2004; Keller 2004).

The aim of this paper is to empirically test these theoretical arguments in a common empirical setting. The baseline question is: what is the role of intensity and variety of internal and external knowledge for regional innovation? This is indeed not a novel question, but the contribution of the paper is to demonstrate the role of knowledge for the regional innovation while bringing together various streams of literature, i.e. the literature concerning regional knowledge production function (RKPF) and the literature concerning international technology diffusion. Furthermore, the paper follows recent literature (Frenken et al. 2007) and distinguishes between related and unrelated variety, making it possible to empirically assess the influence of each type of variety on regional innovation. To the best of the authors' knowledge, this is the first attempt to implement the idea of related and

unrelated knowledge variety in a study of regional innovation.[3]

Using panel data on patent applications distributed across 81 Swedish functional regions, the analyses in the paper show that both the intensity and variety of internal and external knowledge matter for the regional innovation, where related variety seems to play a superior role compared to unrelated variety.

The rest of the paper is organized as follows. Section 2 develops five hypotheses concerning the role of knowledge on regional innovation. Section 3 presents the dataset and the methodology followed in the empirical analysis. Section 4 presents the main findings of the paper. Section 5 summarizes and concludes.

2 The role of intensity and variety of regional knowledge on the regional innovation

A number of studies, both in the field of regional economics and strategic management, have recognized that sustained competitive advantages of regions are related to the ability of regional firms to develop and maintain their innovation (Grant 1996; Krugman 1991; Maskell and Malmberg 1999; Saxenian 1996). Such innovation is furthermore argued to be dependent on the 'intensity' and 'variety' of the knowledge sources available for a region (Audretsch and Vivarelli 1996; Audretsch and Feldman 2004; Beaudry and Schiffauerova 2009).

One source of knowledge is "internal" knowledge sources generated within a region. A large literature on the knowledge production function (KPF) show that new knowledge is essentially generated via "intensity" of R&D activities (Griliches 1979) carried out by firms, university, and research centres (Audretsch and Feldman 1996; Acs et al. 2002). The original Griliches' firm-level KPF framework has been translated to the regional level, so-called regional knowledge production function (RKPF) (Jaffe 1989; Feldman and Florida 1994). This literature emphasizes that the

---

[2] One could ask even a further question: given the importance of knowledge variety (diversity), how to enhance it in the regions/cities (Audretsch and Belitski 2013)?

[3] There are previous studies showing the positive effect of the concept of related industries on regional innovation measures (Feldman 1994; Feldman and Audretsch 1999; Ejermo 2005). Yet, this paper employs the entropy measure, which turns out to be an attractive measure, since it distinguishes between related and unrelated variety.

relevant knowledge for many local firms is the knowledge that 'spills over' from local R&D activities.

In addition, it is shown that the accumulation (intensity) of knowledge per se is not sufficient for a strong innovative performance. The "variety" of knowledge inside a region also matters (Jacobs 1969; Saviotti 1996). Knowledge variety refers to that the knowledge, know-how, and expertise in a region is most often heterogeneous. Exposure to heterogeneous knowledge should improve both the creative potential of firms in the region as well as their ability to develop innovation (Rodan and Galunic 2004). This associates to Schumpeter's (1934) idea of "novelty by combination". Duranton and Puga (2001) called the regions with available variety of knowledge as "nursery cities", because these regions allow firms to try a variety of processes before finding their ideal process innovation, without costly relocation after each trial. As a consequence, the existence of variety of regions' internal knowledge can be considered as an important factor to explain a region's innovation. In this argument, the source of knowledge is external to the firms but still internal to the region, highlighting the role of region (location) on innovation (Feldman 2003).[4] Drawing on above, the innovation of a region rises with both the intensity and the variety of the internal knowledge within the region (Asheim et al. 2011; Berliant and Fujita 2011). This leads to formulate the following two hypotheses:

Hp1 The higher the intensity of the region's internal knowledge, the higher will be the region's innovation.

Hp2 The higher the variety of the region's internal knowledge, the higher will be the region's innovation.

It is also expected that such variety of knowledge can have even more positive impact on innovation, if it is a "related" variety rather than "unrelated" variety. The notion of related variety aims to capture the balance between cognitive proximity and distance across sectors in a region that is needed for knowledge to spill over effectively between sectors (Nooteboom 2000). The unrelated variety measures the extent to which a region is diversified in very different types of activity. According to Frenken et al. (2007), the higher the number of technologically related sectors in a region, the higher

inter-sectorial knowledge-spillovers between those related sectors, and presumably the more learning opportunities for them. This will eventually enhance regional innovation (Feldman 1994). The importance of knowledge-spillovers for regional innovation is illustrated by the following statement in Audretsch and Feldman (2004, p. 2719): "innovative activity should take place in those regions where the direct knowledge-generating inputs are the greatest (e.g. R&D investment), *and* where knowledge spillovers are the most prevalent (which can be achieved by higher related variety in a region, as noted above)". The benefit of related variety of knowledge for innovation-related measures is shown both in the firm-level (Breschi et al. 2003) and in regional-level studies (Feldman 1994; Feldman and Audretsch 1999; Ejermo 2005; Antonietti and Cainelli 2011). In Swedish case, Ejermo (2005) finds a positive effect of "weighted-average-relatedness of neighbours" (WARN) on the patent application. Such a WARN index gets higher value if the "relatedness" between the patenting activities in a region gets higher. The unrelated variety effect, on the other hand, captures the portfolio-effect, which functions as a regional shock absorber (Essletzbichler 2007). That is, when a region has a large number of unrelated industries, it may not be too vulnerable to sector-specific shocks, such as unemployment (Boschma et al. 2012). While unrelated variety may be seen as a variety in general and hence beneficial for innovation, yet, related variety is expected to be more important. The main argument here is that the more related industries in a region, the more possibility of intra-regional knowledge spillover between these related industries, which, consequently, may increase the chance of knowledge generation and hence increasing the regional innovation. This leads to the following additional hypothesis:

Hp3 The impact of related variety on the region's innovation is higher than the impact of unrelated variety.

Processes of knowledge generation and combination are not only activated by means of interactions between local knowledge resources. Knowledge can also be brought into a region from "outside" through the international trade networks. Here, the knowledge sources are external both to the firm and to the region, i.e. global pipeline (Bathelt et al. 2004).[5] This is in line

---

[4] One way to enhance the variety (heterogeneity) of knowledge in the region/city is recently suggested to be the existence of creative class and ethnical diversity, which may enhance knowledge spillover and entrepreneurial opportunity in the region (Audretsch and Belitski 2013).

[5] This has been recognized as one of the stylized fact in the geography of innovation (Feldman and Kogler 2010).

with the literature on international technology diffusion, which proposes international trade as a conduit for flow of knowledge into the local firms within the regions (Keller 2004). This is also in line with 'Learning-By-Exporting' literature, arguing that firms that trade internationally have better access to knowledge about customer preferences, production techniques and foreign technology, which in turn may stimulate innovation and productivity (Clerides et al. 1998; Castellani 2002). All these literature emphasize on the importance of the intensity of the external knowledge as the source of knowledge generation (and consequently innovation capability) within the regions. In addition, there are recent evidence suggesting that the more "related" import and export portfolio of the region, the more learning opportunity within the region (Boschma and Iammarino 2009; Boschma et al. 2012), which in turn can imply for more innovation of the region. Such arguments lead to the following two hypotheses:

Hp4 The higher the intensity of the external knowledge brought into the region, the higher will be the region's innovation.

Hp5 The higher the related external knowledge brought into the region (the more imports and exports are related), the higher will be the region's innovation.

To test the five proposed hypotheses, an econometric analysis on 81 functional regions in Sweden over 2002–2007, controlling for a set of control variables, is applied. This is presented in the rest of the paper.

3 Empirical analysis

3.1 Data

The geographic unit in the empirical analysis is functional regions. The Swedish Development Agency (NUTEK) has divided Sweden into 81 functional regions, each composed of several municipalities. The basic criteria for such division have been the common local labour market (LLM) and commuting time (NUTEK 2005). Andersson and Karlsson (2007) find that knowledge flows in Sweden transcend municipal borders, but they tend to be bound within functional regions. This is because of the fact that functional regions differ from each other in terms of their production of and access to knowledge (Karlsson and Johansson 2006). This makes it plausible to choose functional regions as the unit for the analysis of

innovation, as this level of aggregation should mean that a large part of spatial dependence is internalized within the unit of analysis. Moreover, as it will be shown in Fig. 1, patent applications in Sweden are strongly concentrated in the "islands of innovation" rather than neighbouring regions. This adds additional support that spatial dependency should not be a major concern in this study. Appendix 1 shows the map of the Swedish functional regions.

Two sources of statistics are used to build the dataset: Statistic Sweden (SCB) and the European Patent Office (EPO) database. Data on R&D investment, total employment in two-digit and five-digit NACE industry, higher educated employment, import and export, living place of inventors, and population during 2002–2007 all originate from Statistics Sweden. The European Patent Office (EPO) database provides patent data for Sweden which covers the period 2002–2007.[6] It accounts up to 85 % of all Swedish patent applications in this period. A Swedish patent application is the one that has at least one inventor with living address in Sweden. Patents are regionalized according to the place of the inventors. If a patent application has more than one inventor, following Jaffe et al. (1993), it is equally fractionalized based on the number of inventors. For instance, if a patent application has four inventors, each inventor (and the corresponding functional region that s/he lives) receives 25 % of that patent application. The final dataset is the result of merging the data concerning the determinants of regional innovation with patent data, which provides a balanced panel dataset of 486 observations consisting 81 units (functional regions) over the 6 years period (2002–2007).

3.2 The model and measurements

Analysing the determinants of regional patent (or other measure of innovation like product announcement) is extensively performed in so-called Regional Knowledge Production Function (RKPF) framework (Jaffe 1989; Feldman 1994; Feldman and Florida 1994; Acs et al. 2002). This framework shifts the unit of analysis from traditional firm-level to regional-level, while maintaining the original Cobb–Douglas specification

---

[6] It is preferable to use the EPO data rather than data from Swedish Patent Office, since in recent years the number of Swedish patent applications in EPO is increasing relative to Swedish Patent Office. Hence, it is assumed that EPO data can provide richer dataset in the study period of this paper.

(Audretsch and Feldman 2004). The general specification of RKPF framework is:

$$y_r = \alpha \mathbf{X}_r^{\beta} \qquad (1)$$

where y_r is inventive or innovative output in region r, and $\mathbf{X}_r$ is the vector of innovation inputs within the region r, such as R&D investments, inter-industry knowledge spillover within a region (usually measured by the concentration of related manufacturing industries in the region), and human capital (described in *Hp1*, *Hp2*, *Hp3*).[7] Therefore, the main innovation inputs in RKPF framework have been considered to be "internal" sources of knowledge generated *within* the region. This paper indeed extends the RKPF specification by adding the "external" sources of knowledge brought into the region, as additional inputs (explanatory variables) of innovation (output) of regions. This is motivated by international technology diffusion theory and also so-called Learning-By-Exporting literature (described in *Hp4*, *Hp5*). Such extended specification will be presented in this section, after the proper estimator is chosen.

The negative binomial regression model is applied in order to estimate the relationship between regional innovation, proxied by patent applications, and its determinants presented in Sect. 2. The reason for choosing such estimator is because of the special feature of the dependent variable. The dependent variable *patent application* is count data.[8] It also suffers from over-dispersion as the sample variance is 273 times the sample mean.[9] In order to handle this situation, the

literature suggests several models such as negative binomial, zero-inflated negative binomial (ZINB), and hurdle models (Cameron and Trivedi 2008).[10] The dependent variable does not have many zero values. Only 24 % of the patent values are zero in the sample (119 out of 486 observations), which provides little justification for using a ZINB model.[11] Indeed, a Vuong test of zero-inflated negative binomial versus (standard) negative binomial speaks in favour of (standard) negative binomial. The hurdle model is not a preferred model, too, for the same reason as for the zero-inflated models. Using the negative binomial model, the formulation for the regional innovation, which is an extension of RKPF, is written as follows:

$$\Pr(\mathbf{y_{rt}} = \widetilde{\mathbf{y_{rt}}} | \mathbf{X_{1rt} X_{2rt}}\ \mathbf{Z_{rt}}, \sigma_{rt}) = \frac{\mathbf{e}^{-\lambda_{rt}} \times (\lambda_{rt})^{\widetilde{\mathbf{y_{rt}}}}}{\widetilde{\mathbf{y_{rt}}}!} \qquad (2)$$
$$\widetilde{\mathbf{y_{rt}}} = \mathbf{0, 1, 2, 3, \ldots}$$

where,

$$\lambda_{rt} = \exp(\beta_1 X_{1rt} + \beta_2 X_{2rt} + \beta_3 Z_{rt}) \times \exp(\sigma_{rt})$$

where y_{rt} is the number of patent applications in functional region r in year t, $\mathbf{X}_{1rt}$ is the vector of internal knowledge variables, $\mathbf{X}_{2rt}$ is the vector of external knowledge variables, $\mathbf{Z}_{rt}$ is the vector of control variables, β_j are the coefficient parameters to be estimated, and $\exp(\sigma_{rt})$ is assumed to have a gamma distribution with mean 1 and variance *alpha*, which can be estimated from the data. Alpha is the over-dispersion parameter, which corrects for the over-dispersion by adjusting the variance independently from the mean (Cameron and Trivedi 2008).

---

[7] The unit of analysis in RKPF has been either region-technology (Jaffe 1989; Feldman and Florida 1994; Ponds et al. 2010) or region (ACS et al. 2002; Ejermo 2005; Fritsch and Slavtchev 2007). This paper chooses the later alternative, while controlling for industry heterogeneity across regions.

[8] It is worthy to note that, however, the patent data is a non-integer data, since the patent data is fractionalized (based on the number of inventors belonging to different functional regions). This could violate the usage of negative binomial regression, since this technique is designed for count data (integer data). To avoid this possible violence, the rounded value of fractionalized patent data is used in the regression. There are two groups of data that are in the risk of being under/overvalued after rounding: (1) the observations with the patent value between 0 and 0.5 and (2) with the value between 0.5 and 1. Nonetheless, the numbers of observation in former group is only ten and in the latter one is only eleven. More importantly, the result of binomial regression before and after rounding is quite similar.

[9] The mean value for patent application is 24 and the variance is 6,560.

[10] Since the mean and variance are not equal, therefore the equidispersion assumption is violated, which implies that the estimations based on Poisson and Zero-inflated Poisson models are not the preferred options.

[11] Even if there would be many zero value in the data, it does not necessarily mean that zero-inflated models can be the best option (Cameron and Trivedi 2008, p. 605), since it must be possible to distinguish between 'true zeros' and 'excess zeros' in order to be reasonable to use zero-inflated models. The mechanism for distinguishing these two types of zero is not clear in the patent application data, hence the use of zero-inflated models seems to be implausible. An example of the a situation where it is possible to distinguish between true zeros and excess zero is when a researcher wants to explain the amount of cigarettes smoked per day, while s/he has a survey containing both smokers (can cause true zeros) and non-smokers (causing excess zeros) (Cameron and Trivedi 2008, p. 584).

The Likelihood-Ratio (LR) test of panel versus pooled has been always in favour of panel models (reported in Table 2), hence the panel application of negative binomial model is chosen. For the dependent variable and most of the regressors, the vast majority of the variation in the data consists of the between-variation rather than the within-variation. Therefore, the fixed-effects estimator may not be very efficient, since it relies on within-variation (Andersson and Lööf 2011). Furthermore, it is argued that fixed-effects estimator may even wrongfully include the impact of those variables which exhibit only slight changes over time, i.e. in this paper Related and Unrelated Variety (Fritsch and Slavtchev 2007). Therefore random-effects estimator (RE) is used in the panel models.[12]

3.2.1 The dependent variable

The phenomenon under study is the innovation of Swedish functional regions. The number of patent applications for 81 Swedish functional regions during 2002–2007 is used as a proxy for it (see Jaffe and Trajtenberg 2005; Acs et al. 2002). Patents have been found to be a good proxy of innovative activity in general (Griliches 1990) and for regional-level analysis in particular (Acs et al. 2002).[13] This is because patents are granted for inventions which are novel, inventive, and have industrial application (Andersson and Lööf 2011).[14]

3.2.2 Independent variables

Based on the underlying theories discussed in part 2, the independent variables in this paper are grouped into two categories: (1) internal knowledge and (2) external knowledge.

Internal knowledge: To capture the characteristics of the region's internal knowledge, in terms of intensity and variety, three different variables are included: the intensity of R&D activities of the region, the unrelated variety of knowledge within the region, and the related variety of knowledge within the region. The first one captures intensity aspect and second and third variables capture variety aspect.

The measure of the intensity of R&D activity ($R\&D_{rt}$) is the log amount of R&D investments of the region r in year t in Million Swedish Kroner (MSEK).[15]

The unrelated variety (URV) of knowledge within the region is measured as the entropy at the two-digit level (Frenken et al. 2007; Boschma et al. 2012). The unrelated variety (URV) index for region r in year t is given by:

$$URV_{rt} = \sum_{g=1}^{G} Pg_{rt} \log_2 \left(\frac{1}{Pg_{rt}} \right) \qquad (3)$$

where, Pg_{rt} is the employment share in 2-digit industry NACE code for region r in year t and G is the maximum number of two-digit industries in region r and year t. Theoretically, URV_{rt} gets the minimum value of 0 and maximum value of $\log_2 G$. The minimum value happens when all employees in the region working in the same 2-digit industry, hence no variety exists at all. The Maximum value happens when there is an equall distribution of employees over all 2-digit industries. In this case G encompasses all 2-digit sectors of economy.[16]

For measuring the related variety, following Frenken et al. (2007), it is assumed that five-digit industries are technologically related when they share the same two-digit class. These industries are perceived to show some degree of cognitive proximity, because these five-digit sectors (e.g., sub-branches in chemicals) will share some

---

[12] Here 'random effects' apply to the distribution of the over-dispersion parameter (*alpha*), which is the same for all observations in the same group (here functional region) but varies randomly from group to group.

[13] Acs et al. (2002) compared the number of new products and patents across US regions and conclude; "the empirical evidence suggests that patents provide a fairly reliable measure of innovative activity" (p. 1080).

[14] It must be acknowledge that the use of patents as indicators for innovation is not undisputed. It is argued that not all patents are innovations and not all innovations are patented (Griliches 1990).

[15] The R&D investment refers to corporate R&D investment. In Swedish case it is documented that the corporate R&D (not university R&D) is the one that has the significant impact on innovation of the regions (Gråsjö 2006). Gråsjö (2006) used patent application as the measure of innovation of Swedish regions. He found significant impact of corporate R&D (but not university R&D) on innovation. The lack of significant impact of university R&D has been also found in explaining the Swedish export (Gråsjö 2006) as well as Swedish firm formation (Karlsson and Nyström 2011).

[16] In addition to Unrelated Variety, this paper initially considered the measured of the pure variety, which is a disentangled measure of variety in five-digit sectors within the regions (Boschma et al. 2012). Substituting the unrelated variety with pure variety measures reveals very similar results.

technology and product characteristics in the same two-digit class (e.g., chemicals). At the same time, these industries are considered to show some degree of cognitive distance, because these sectors differ at the five-digit level. Then, the more sectors at the five-digit level within each two-digit level in a region, higher the value of related variety. Therefore, related variety (RV) in region r and time t, as the weighted sum of entropy within each two-digit sector, is given by:

$$RV_{rt} = \sum_{g=1}^{G} Pg_{rt}Hg_{rt} \qquad (4)$$

where:

$$Hg_{rt} = \sum_{i \in Sg} \frac{Pi_{rt}}{Pg_{rt}} \log_2\left(\frac{1}{Pi_{rt}/Pg_{rt}}\right)$$

where, Pi_{rt} is the employment share in 5-digit NACE code for region r in year t, Pg_{rt} is the employment share in 2-digit NACE code for region r in year t, G is the maximum number of two-digit sectors in region r and year t, I is the maximum number of five-digit in the region, and all five-digit sectors i fall exclusively under a two-digit sector Sg. Similiar to URV_{rt}, RV_{rt} gets the theoretical minimum value of 0 when all employees of the region in 2-digit industry g are working within the same the same 5-digit industry i, where i ∈ Sg. The maximum value, $\log_2 I$, is achived when there is an equal distribution of employees over all 5-digit industries i, where i ∈ Sg.

External knowledge: To capture the characteristics of the region's external knowledge, in terms of intensity and variety, two different variables are included: the amount of export and import, as intensity measure, and related trade variety index, as variety measure.

The measure adopted as a proxy for the intensity of external knowledge brought into the region relates to the amount of the international trade linkages of each functional region (Boschma and Iammarino 2009). It is measured as:

$$IMP\_EXP_{rt} = Ln(Import\ value_{rt} + Export\ value_{rt}) \qquad (5)$$

where, *Import value*$_{rt}$ and *Export value*$_{rt}$ are the value of the import and export in manufacturing in region r in year t, respectively. The higher the value of IMP_EXP_{rt}, the greater is the external knowledge that flows into the region.

The measure adopted as a proxy for the variety of external knowledge brought into the region is the Trade Related Variety (Boschma et al. 2012). It aims to measure the extent to which the export portfolio of a region is related to its import portfolio. Let 1 be a five-digit industry within the two-digit class $I(i)$, with $i = 1,\dots, n$. Then following Boschma and Iammarino (2009), trade related variety (TRV) is given as follows:

$$TRV_{rt} = \sum_{i} Import_5^M(i)Export_5(i) \qquad (6)$$

where, $Import_5^M(i)$ is the import entropy in five-digit industries other than *1*, but within the same two-digit industry $I(i)$, i.e. $(i \neq 1)$. $Export_5(i)$ is the relative size of the five-digit export industry *1* (with $i = 1,\dots, n$) in the entire provincial export.

3.2.3 Control variables

Several control variables are considered: population density, human capital, manufacturing concentration index, high-technology manufacturing concentration index, number of high-tech large manufacturing firms, and year dummies. Population density (*POPULATION*) controls for the size of the regions and captures urbanization economies. It is measured as the population per square kilometre in each region each year. Population density is expected to have positive effect on innovation (Feldman 1994).

Human capital (HC) is a standard variable in KPF framework. It is shown to have the significant positive impact on innovation in firm level (Andersson and Ejermo 2005) and regional level studies (Lee et al. 2010). Such a positive effect on innovation is motivated generally by endogenous growth theory (Romer 1986), and specifically by Lucas (1988)'s model, arguing that the ability to develop new technology depends on the average level of human capital in the local economy. It is measured as the fraction of higher educated employees (employees with three or more years of university education) in region r year t.

It is shown that sectors differ in their propensity to patent. First of all, service sector is less likely to patent its knowledge production compared with manufacturing (Hipp and Grupp 2005). In order to incorporate this argument, the paper includes location quotient of the manufacturing specialization (LQ_MAN). It is calculated as follows:

$$LQ\_MAN_{rt} = \frac{\left(\frac{\text{Manufacturing employment in region } r \text{ year } t}{\text{Total employment in region } r \text{ year } t}\right)}{\left(\frac{\text{Manufacturing employment in economy year } t}{\text{Total employment in economy year } t}\right)}$$

$$(7)$$

The higher the value of LQ_MAN, the higher concentration of manufacturing sectors in the region. This variable controls for the fact that manufacturing sectors have higher propensity to patent than service sectors, and expected to have positive sign (Fritsch and Slavtchev 2007; Paci and Usai 1999).

Second, even within manufacturing, sectors have shown different behaviour in terms of propensity to patent, because different sectors have different technology and innovation opportunities, and are thus characterized by different technological regimes (Malerba and Orsenigo 1997). For instance pharmaceutical and chemical sectors are more likely to patent because they are in high-tech manufacturing sectors (Scherer 1983). As for controlling this second point, the paper includes location quotient of the High-Tech manufacturing sectors (LQ_HT).[17] It is calculated as follows:

$$LQ\_HT_{rt}$$
$$= \frac{\left(\frac{\text{High}-\text{Tech manufacturing employment in region } r \text{ year } t}{\text{Total employment in region } r \text{ year } t}\right)}{\left(\frac{\text{High}-\text{Tech manufacturing employment in economy year } t}{\text{Total employment in economy year } t}\right)}$$

$$(8)$$

This variable is also expected to have positive effect on innovative capability.[18] Both LQ_MAN and LQ_HT are controlling for heterogeneous industry structure across regions, which stem from the heterogeneous propensity of sectors to patent.

One would like to see if the result is driven merely by the presence of a few large firms in industries prone to patenting activity or not. In order to account for this, the number of large firms (firms with 500 or more employees) in High-Tech manufacturing sectors in the region (LRG_HT) is included as another control

variable. It is expected that the presence and dominance of large firms would have negative effect on regional innovative capability, ceteris paribus (Acs et al. 1994, 2002).

Finally, year dummies are included to capture heterogeneity between years. This will take into account the macroeconomic effects and business cycles that may affect regional patent applications. In order to reduce simultaneity concerns, 1 year lag for all right hand side variables is used in the subsequent empirical analysis.[19] The definitions of all variables are documented in Appendix 2.

3.3 Data descriptions and correlations

The distributions of main variables, i.e. PATENT, R&D, RV, URV, EXP_IMP, TRV, over functional regions are illustrated in Fig. 1. The values are the average value during 2002–2007.

The patent applications are geographically concentrated in Sweden, which is well in line with previous findings in US and Europian studies (Audretsch and Feldman 1996; Breschi 1999). In other words, it also shows the clear evidence of over-dispersion of patent application (Bettencourt et al. 2007), with four regions over performing than other regions, i.e. Stockholm, Gothenburg, Malmö, and Linköping, in which the first three ones are the three Swedish metropolitan areas. On the other extreme, five functional regions do not have any patent application during the study period, i.e. region 66, 67, 73, 76, and 78 (see Appendix 1 for the name of regions). One interesting point is that RV is the highest exactly among those four regions with highest patent applications, while this is not the case for URV. This is already an initial indication of superior role of RV compared with URV. Another interesting point is the regional portfolio of Stockholm, as the highest producer of patent applications: Stockholm develops its variety more in a related sense, rather than unrelated sense.

The correlation matrix and descriptive statistics for all variables are presented in Table 1.

---

[17] High-Tech manufacturing sector is defined based on OECD classifications. It consists of following NACE codes: 2433, 30, 32, 33 and 353. Similar classification is used in other patent studies, for instance in Andersson and Lööf (2011).

[18] The value of LQ_HT is normalized by $LQ\_HT^{norm} = (LQ_{HT} - 1)/(LQ_{HT} + 1)$. LQ_HT^{norm} is systematically distributed between -1 and $+1$ (Fritsch and Slavtchev 2007; Paci and Usai 1999). The same normalization is done for LQ_MAN.

[19] It should be acknowledge that one year may not be the best lagging option, as other studies uses 2 or 3 years lag (Fritsch and Slavtchev 2007; Ponds et al. 2010), although there are indeed one-year-lag studies, too (Crescenzi et al. 2012). However, the sample would have been substantially small if the lag was increased.

Fig. 1 Distribution of main variables over Swedish functional regions (average value during 2002–2007)
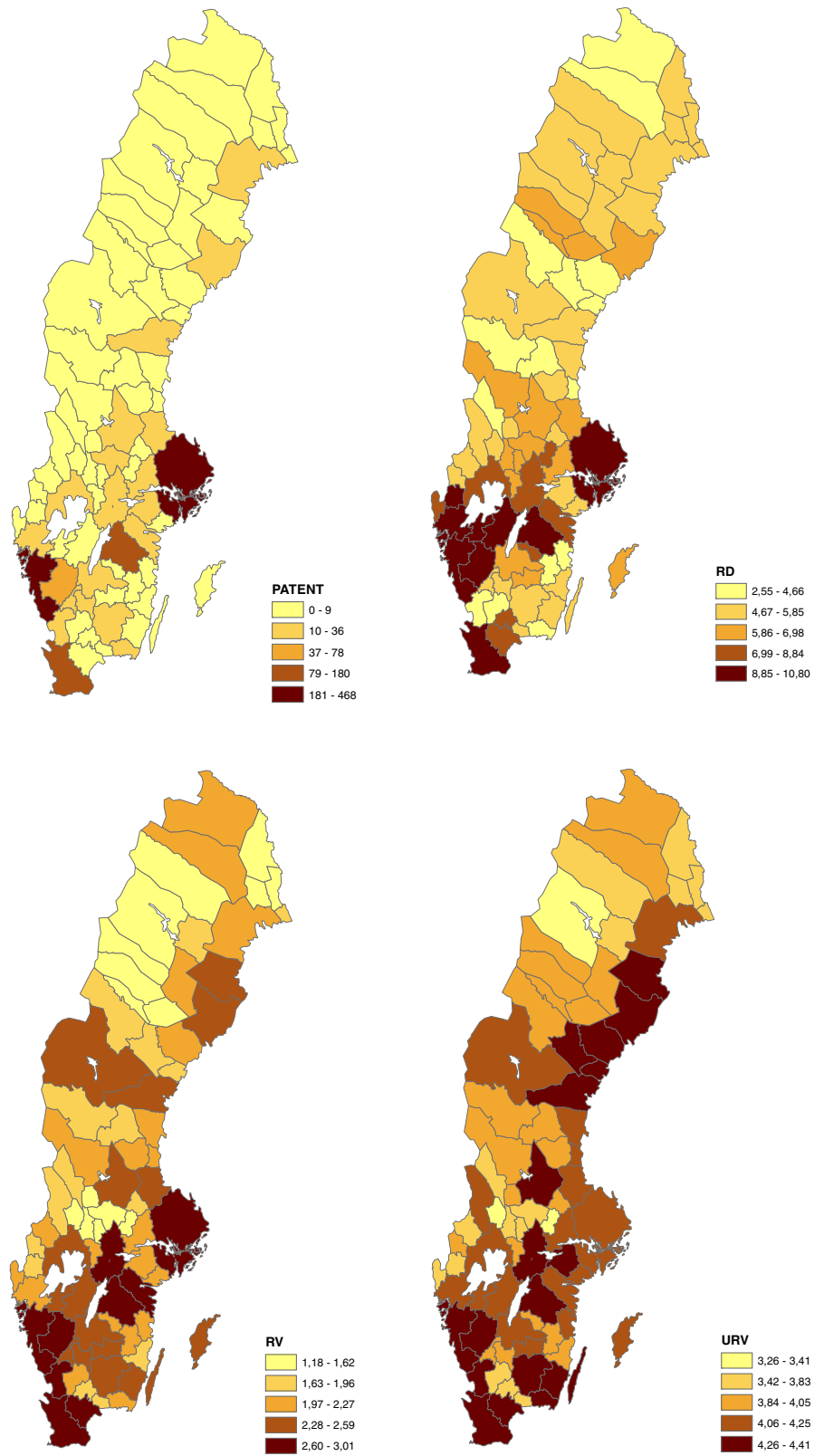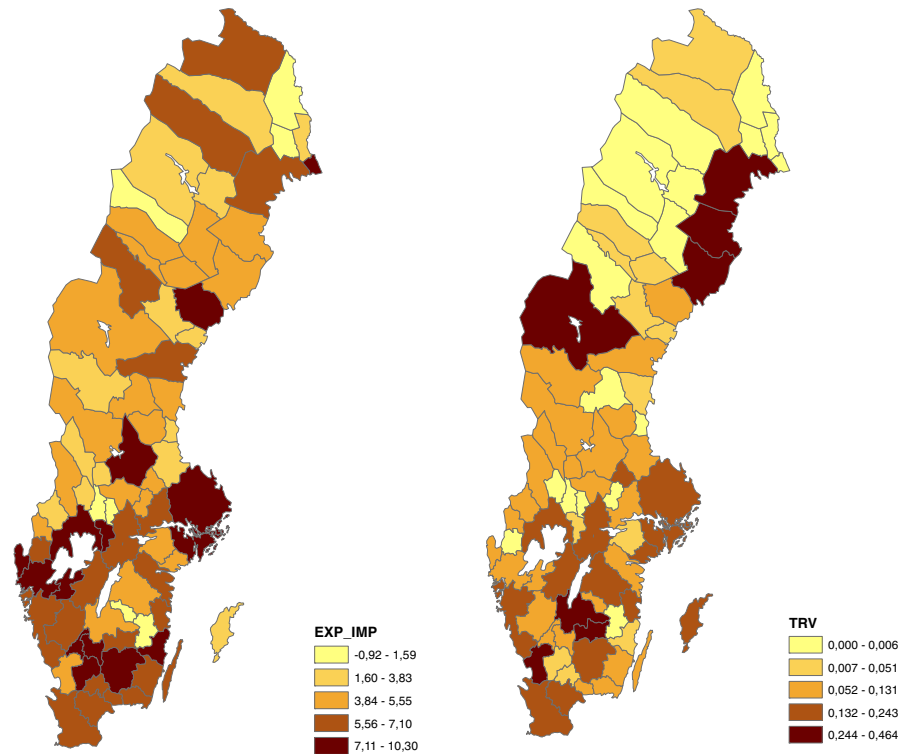
Fig. 1 continued

Table 1 Correlation matrix and descriptive statistics

Variables	PATENT	R&D	URV	RV	IMP_EXP	TRV	HC	POP	LQ_MAN	LQ_HT	LRG_HT
PATENT	1										
R&D	0.488	1									
URV	0.228	0.192	1								
RV	0.435	0.403	0.678	1							
IMP_EXP	0.455	0.361	0.484	0.654	1						
TRV	0.226	0.174	0.359	0.461	0.346	1					
HC	0.545	0.449	0.479	0.690	0.628	0.404	1				
POP	0.886	0.535	0.380	0.607	0.553	0.279	0.671	1			
LQ_MAN	0.294	0.002	0.117	−0.21	−0.09	−0.12	−0.06	−0.16	1		
LQ_HT	0.261	0.171	0.384	0.213	0.143	0.138	0.302	0.292	0.038	1	
LRG_HT	0.818	0.373	0.142	0.281	0.351	0.209	0.406	0.293	−0.264	0.24	1
Observations	486	486	486	486	486	486	486	486	486	486	486
Mean	24.02	6.15	4.03	2.12	16.27	0.10	0.03	28.24	0.09	−0.39	0.24
SD	80.99	1.95	0.29	0.45	3.23	0.13	0.02	39.53	0.17	0.44	1.21
Min	0	2.04	2.76	1.04	0	0	0.007	0.24	−0.77	−1	0
Max	632	11.12	4.46	3.13	24.09	0.52	0.125	250.09	0.40	0.75	14

Not only internal but also external knowledge sources (both intensity and variety aspects) are positively correlated with patent application. Interestingly RV is correlated with patent application twice as the

URV with patent application. This is again an indication for the superior role of related variety compared with unrelated variety, as noted in Fig. 1. The variance inflation factor (VIF) test is performed to check for

Table 2 Determinants of patent application (2002–2007); panel negative binomial estimation

VARIABLES	(1)	(2)	(3)	(4)	(5)
R&D	0.179***	0.157***	0.113**	0.131***	0.080*
(lagged 1 year) (log)	(0.049)	(0.048)	(0.045)	(0.046)	(0.041)
URV	0.909**	0.686*	0.900**	0.261	0.633
(lagged 1 year)	(0.383)	(0.402)	(0.389)	(0.420)	(0.403)
RV	1.039***	1.030***	0.675**	1.413***	0.993***
(lagged 1 year)	(0.272)	(0.272)	(0.275)	(0.285)	(0.280)
HC	1.440***	1.481***	1.271***	1.518***	1.336***
(lagged 1 year)	(0.209)	(0.211)	(0.205)	(0.203)	(0.197)
IMP_EXP		0.044***	0.040**	0.038**	0.030*
(lagged 1 year) (log)		(0.016)	(0.016)	(0.016)	(0.016)
TRV		−0.211*	−0.218*	−0.200	−0.195
(lagged 1 year)		(0.127)	(0.120)	(0.131)	(0.126)
Population			0.009***		0.010***
(lagged 1 year) (log)			(0.003)		(0.002)
LQ_MAN				1.641***	2.121***
(lagged 1 year)				(0.570)	(0.540)
LQ_HT				0.376**	0.259
(lagged 1 year)				(0.182)	(0.170)
LRG_HT					−0.006
(lagged 1 year)					(0.011)
LR test of Alpha = 0	1,117.86	1,026.34	605.77	1,019.33	448.27
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LR test versus pooled	328.65	318.99	197.59	319.16	158.91
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	405	405	405	405	405
Nr of Regions	81	81	81	81	81
Year Dummy	Yes	Yes	Yes	Yes	Yes
AIC	1,922.45	1,891.28	1,882.08	1,883.07	1,869.89

Dependent variable in all models: Number of patent applications in Swedish functional regions over 2002–2007

The table reports coefficient parameters with standard errors in parentheses

For LR (Likelihood Ratio) test versus pooled, Prob $\geq \chi^2$ in parentheses

For LR test of Alpha = 0, Prob $\geq \chi^2$ in parentheses. Alpha: Over-dispersion parameter estimated with pooled models

AIC akaike information criterion

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

multicollinearity between independent variables. All independent variables got the value of lower than 4 in this test and the overall VIF score was equal to 2.16. Therefore, it is expected that multicollinearity does not substantially bias the regression results in part 4.[20]

---

[20] The VIF test is performed after the conventional OLS regressions. There is no formal threshold for variance inflation factor test, but as a rule of thumb the VIF score below 10 (or sometimes 5) is said to be the evidence of quite mild multicollinearity.

4 Empirical results

The results of negative binomial random effect estimation of knowledge-based determinants of patent application for 81 Swedish functional regions over the period of 2002–2007 are reported in the Table 2.

First model (column (1)) considers only the effect of the intensity and variety of the region's "internal" knowledge. Second model adds the intensity and variety of the "external" knowledge. Third model controls for

the size of regions by adding POPULATION. Fourth model takes into account the industry heterogeneity across regions by adding LQ_MAN and LQ_HT. Finally, fifth model add LRG_HT to take into account the possible dominancy of few (large) high-tech firms in the region. This model is the full model which includes all explanatory variables and control variables.

R&D investment and HC are positive and highly significant in all models, as expected. They show the importance of intensity of internal knowledge (i.e. generated inside the region) for innovation of the region. Human capital in particular signifies the importance of highly educated individuals for producing patents, and that there are positive externalities to schooling. URV is always positive showing the importance of variety (in general sense) for innovation. However, its significance diminishes from model 1 to the model 5 (full model). On the other hand, the interesting point is that the related variety of knowledge within the regions (RV) is always positive and significant, even after controlling for population (in model (3)), heterogeneous industry structure across regions (in model (4)), and firms size composition of regions (in model (5)). The robust result concerning RV can be explained as in Hypothesis 3: A region with higher RV can enjoy the higher learning opportunity and knowledge spill-overs between the existing related sectors within that region (compared with regions with high URV) (Frenken et al. 2007), which eventually lead to higher innovation for the region (Feldman 1994; Audretsch and Feldman 2004). These results confirm hypotheses 1, 2, and 3.

The intensity of external knowledge (IMP_EXP) is positive and significant in the all models in which it is included, thus confirming hypothesis 4. However, unlike the expectation, the TRV measure shows negative sign, nevertheless, the significance is weak and it is not a robust result to explain the innovation of the regions. This means the null of hypothesis 5 cannot be rejected. TRV indeed has shown some vague results in previous studies: while it shows the positive and significant effect on regional employment growth, it shows no significant impact on regional value-added growth and labour-productivity growth, even with the negative sign in the later one (Boschma and Iammarino 2009).

As for control variables, population density is always significant and positive, which is in line with previous research on RKPF framework (Feldman 1994). This shows the positive effect of scale or (pure) urbanization economies. Model (4) shows that regions with

concentration of manufacturing sectors (LQ_MAN) in general and high-tech manufacturing sectors (LQ_HT) in particular are performing better in terms of applying for patent, as expected (Scherer 1983; Hipp and Grupp 2005). Moreover, controlling for such industry heterogeneity (i.e. acknowledging different propensity to patent in different industries) across regions did not changed the main results concerning internal and external knowledge. Finally in the full model, LRG_HT shows the negative sign. Hence it seems merely a few large firms prone to patenting activity are not driving the innovation of regions; rather the patenting activities seem to be also spread out between many smaller firms in the region.[21] This is in line with previous studies suggesting that regions dominated by small firms have higher innovation, ceteris paribus (Acs et al. 1994, 2002). However, the coefficient is not significant, makes one to be cautious about interpretation. Nevertheless, this variable controls for the possible concentration of patent activities in few (large) firms prone to patenting, and the main results remained the same.

Since the Maximum Likelihood Estimation (MLE) is used, one way to compare the models with each other is to use Akaike information criterion (AIC) or Bayesian information criterion (BIC). Both criteria get smaller when moving from model (1) to (2) (BIC is not reported in Table 2). This means that by adding the external knowledge variables in model (2), this model is getting better in terms of fitness compared with model (1), which only includes internal knowledge, while there is no evidence of over-fitting. In other words, internal knowledge and external knowledge *together* can produce the better fit for modelling the patent application compared with including only one of them. Controlling for POPULATION in model (3) further improved the model. While moving to model (4) did not improve the model, model (5), which is the full model, turns out to be the best model in terms of AIC. The same evidence is also obtained by performing the Likelihood Ratio test of restricted versus unrestricted models, when moving between models.[22] This can be seen as fulfilling the stated aim of the paper, i.e. to empirically test various theoretical conjectures (RKPF and international trade theory) in a common empirical setting.

---

[21] Using the Location Quotient instead of absolute number of High-Tech large manufacturing firms revealed the same result.

[22] LR test of restricted versus unrestricted models is not reported in Table 2 and is available upon request.

The likelihood ratio (LR) test of including alpha, over-dispersion parameter, is reporter in Table 2, too. In all models the null hypothesis of alpha equal to zero in strongly rejected. This means over-dispersion parameter is significantly different from zero and thus shows (again) that the negative binomial is a preferred estimation strategy over the Poisson or zero-inflated Poisson models. Similar estimation choice has been preferred by previous studies using Swedish patent application data in regional-level (Ejermo 2005) and firm-level studies (Andersson and Lööf 2011).

One may argue that the dominant and positive effect of related variety on innovation may not be equally strong in various stages of Industry/Product Life Cycle. Specifically, one may expect the lower effect of related variety (and variety in general) in later stages of ILC, as the later stages are characterized by dominant design, standardization, and less heterogeneity of the firms and products (Vernon 1966; Utterback and Abernathy 1975; Duranton and Puga 2001). Yet, the data at hand does not allow considering the effect of ILC on related variety and other explanatory variables, leaving such interesting point for further research.

5 Conclusion

This paper analysed the effect of (1) intensity and variety aspects of internal knowledge and (2) intensity and variety aspects of external knowledge on the regional innovation, measured by patent applications. As for variety aspect, the paper distinguished between related and unrelated variety for internal knowledge, and used trade related variety to capture the variety of external knowledge.

The results of the empirical analysis show that the innovation of regions rises with both the intensity and the variety of the internal knowledge. The interesting point is that when it comes to variety of internal knowledge within the region, knowledge variety per se does not substantially affect regional innovativeness, as captured by 'unrelated' variety, but it has the robust and positive impact if it is a 'related' variety. This finding is line with previous literature, though this seems to be the first attempt to implement the idea of related and unrelated knowledge variety in a study of regional innovation (see Glaeser et al. 1992; Frenken et al. 2007). The results also suggest that the intensity of external knowledge flowing into a region has a positive effect on innovation of a region, which is in line with international trade diffusion

(Keller 2004) and learning-by-exporting literature (Clerides et al. 1998; Castellani 2002; Andersson and Lööf 2009). The results are not robust concerning the effect of trade related variety, aiming to capture the relatedness of the import and export portfolio of regions. Specifically on internal and external knowledge, these two categories *together* produce a better fit for modelling the patent application compared with including only one of them. In other words, this is fulfilling the stated aim of the paper, i.e. to empirically test various theoretical conjectures (RKPF and international trade theory) in a common empirical setting. Further, some of the control variables included in the analysis provided additional insight.

In a nutshell, some regions produce more patent than others in Sweden because of several reasons: (1) they are better in generating internal knowledge (both variety and intensity aspect), (2) they are benefiting more from external knowledge flowing into the region from outside through international trade linkages (intensity aspect), (3) they are dominated by High-Tech manufacturing sectors, and (4) they are benefiting from urbanization (scale) economies.

What conclusions can be drawn with respect to policy? The main result shows that having related industries within a region enhances the regional innovation (because of knowledge spillover occurring between those related industries). This implies that regions need to be smart and develop the portfolio of complementary sectors, e.g. by having the related variety portfolio. Policy makers may help to this in two ways. First, one might think of a targeted policy for attracting the related sectors to the given region. However, it is shown that the entry and exit of sectors to the regions is governed by a self-selection process of firms (belonging to sectors) and associated path-dependency process, rather than targeted policy (Neffke et al. 2011). Instead, what policymakers may do is to look at the past, identify the potential sectors that could have contributed to the related variety portfolio of a given region (but they never came to the region yet), and eventually remove the possible bottlenecks that have resisted the entry of those related sectors to the region (Neffke et al. 2011). Second, another way of helping a region to have the variety of knowledge is to attract the creative class individuals into a region. Recent studies suggest that creativity has been a missing pillar in the theory of knowledge spillover (Audretsch and Belitski 2013). Diversity, openness to other cultures, and portion of Bohemians all contribute to higher creativity level in

the regions and hence higher knowledge spillover (Boschma and Fritsch 2009; Audretsch and Belitski 2013) and eventually more innovation. In both of these ways, policymakers may help the regions to develop the related variety type of portfolio, however, only in the long run. Such policies to promote the variety in the region can be a complementary one beside the classic policy of increasing the R&D intensity of the regions. That is having both variety and intensity in mind as complementing each other for having a higher regional innovation.
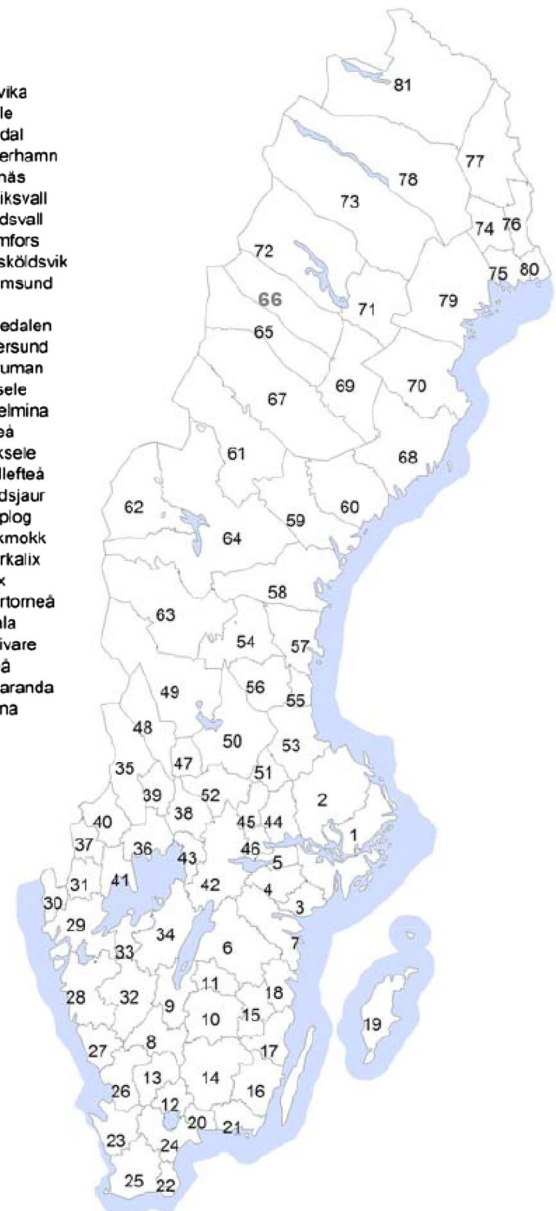
Appendix 1

See Fig. 2.

Fig. 2 Map of Sweden divided to 81 functional regions (local labour market). *Source*: NUTEK (2005)

1	Stockholm	52	Ludvika
2	Uppsala	53	Gävle
3	Nyköping	54	Ljusdal
4	Katrineholm	55	Söderhamn
5	Eskilstuna	56	Bollnäs
6	Linköping	57	Hudiksvall
7	Norrköping	58	Sundsvall
8	Värnamo	59	Kramfors
9	Jönköping	60	Örnsköldsvik
10	Nässjö	61	Strömsund
11	Tranås	62	Åre
12	Älmhult	63	Härjedalen
13	Ljungby	64	Östersund
14	Växjö	65	Storuman
15	Vimmerby	66	Sorsele
16	Kalmar	67	Vilhelmina
17	Oskarshamn	68	Umeå
18	Västervik	69	Lycksele
19	Gotland	70	Skellefteå
20	Karlshamn	71	Arvidsjaur
21	Karlskrona	72	Arjeplog
22	Simrishamn	73	Jokkmokk
23	Helsingborg	74	Överkalix
24	Kristianstad	75	Kalix
25	Malmö	76	Övertorneå
26	Halmstad	77	Pajala
27	Varberg	78	Gällivare
28	Göteborg	79	Luleå
29	Trollhättan	80	Haparanda
30	Strömstad	81	Kiruna
31	Bengtsfors		
32	Herrljunga		
33	Lidköping		
34	Skövde		
35	Torsby		
36	Karlstad		
37	Årjäng		
38	Filipstad		
39	Hagfors		
40	Arvika		
41	Säffle		
42	Örebro		
43	Karlskoga		
44	Västerås		
45	Fagersta		
46	Köping		
47	Vansbro		
48	Malung		
49	Mora		
50	Falun		
51	Avesta		

Appendix 2

See Table 3.

Table 3 Variable definitions and measures

	Variables	Definitions/aims	Measures
Dependent variable	$PATENT_{rt}$	Innovation in region r year t	Number of patent application in region r year t, fractionalized by the inventors' living place
Independent variables			
Internal knowledge			
Intensity	$R\&D_{rt}$	Corporate R&D investment in region r year t	$\ln(R\&D$ investment region r year t in MSEK$)$
Variety	URV_{rt}	Unrelated Variety: Degree of sectoral unrelated variety of region r year t	Entropy measure in two-digit sector
	RV_{rt}	Related Variety: Degree of sectoral related variety of region r year t	Weighted sum of entropy within each two-digit sector
External knowledge			
Intensity	IMP_EXP_{rt}	Amount of import and export in region r year t (in manufacturing)	$\ln(Import\ value\ rt + Export\ value\ rt)$
Variety	TRV_{rt}	Trade Related Varity: Degree of relatedness between export and import portfolio of the region r in year t	Weighted sum of import entropy within each two-digit sector, excluding the same five-digit export
Control variables	POP_{rt}	Population density in region r year t	$\ln\left(\frac{population\ rt}{KM2}\right)$
	HC_{rt}	Human Capital: fraction of higher educated employees (HEE) to the total employees in region r year t	$\frac{HEE\ rt}{Total\ employees\ rt}$
	LQ_MAN_{rt}	Manufacturing concentration index: Share of manufacturing of region r year t to the total economy employment	$\frac{\left(\frac{Manufacturing\ employment\ in\ region\ r\ year\ t}{Total\ employment\ in\ region\ r\ year\ t}\right)}{\left(\frac{Manufacturing\ employment\ in\ economy\ year\ t}{Total\ employment\ in\ economy\ year\ t}\right)}$
	LQ_HT_{rt}	Hi-Tech Manufacturing concentration index: Share of Hi-Tech manufacturing of region r year t to the total economy employment	$\frac{\left(\frac{Hightech\ manufacturing\ employment\ in\ region\ r\ year\ t}{Total\ employment\ in\ region\ r\ year\ t}\right)}{\left(\frac{Hightech\ employment\ in\ economy\ year\ t}{Total\ employment\ in\ economy\ year\ t}\right)}$
	LRG_HT_{rt}	Number of Large High-Tech firms	Number of Large High-Tech manufacturing firms (employee $\geq$ 500 and NACE code: 2,433, 30, 32, 33 and 353) in region r year t
	YEAR DUMMIES	Capturing heterogeneity between years	y2002–y2007

References

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of knowledge. *Research Policy, 31*(7), 1069–1083.

Acs, Z. J., Audretsch, D. B., & Feldman, M. P. (1994). R&D spillovers and recipient firm size. *The Review of Economics and Statistics, 76*(2), 336–340.

Andersson, M., & Ejermo, O. (2005). How does accessibility to knowledge sources affect the innovativeness of corporations? Evidence from Sweden. *The Annals of Regional Science, 39*, 741–765.

Andersson, M., & Karlsson, C. (2007). Knowledge in regional economic growth-the role of knowledge accessibility. *Industry & Innovation, 14*(2), 129–149.

Andersson, M., & Lööf, H. (2009). Learning-by-exporting revisited: The role of intensity and persistence. *The Scandinavian Journal of Economics, 111*(4), 893–916.

Andersson, M., & Lööf, H. (2011). Small business innovation: Firm level evidence from Sweden. *Journal of Technology Transfer*. doi:10.1007/s10961-011-9216-9.

Antonietti, R., & Cainelli, G. (2011). The role of spatial agglomeration in a structural model of innovation, productivity and export. *Annals of Regional Science, 46*(3), 577–600.

Asheim, B. T., Boschma, R., & Cooke, P. (2011). Constructing regional advantage: Platform policies based on related variety and differentiated knowledge bases. *Regional Studies, 45*(7), 893–904.

Audretsch, D. B., & Belitski, M. (2013). The missing pillar: The creativity theory of knowledge spillover entrepreneurship. *Small Business Economics, 41*, 819–836.

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American Economic Review, 86*(3), 630–640.

Audretsch, D. B., & Feldman, M. P. (2004). Knowledge spillovers and the geography of innovation. In J. V. Henderson & J. Thisse (Eds.), *Handbook of urban and regional economics: Cities and geography* (Vol. 4, pp. 2713–2739). Amsterdam: North Holland Publishing.

Audretsch, D. B., & Vivarelli, M. (1996). Small firms and spillovers: Evidence from Italy. *Small Business Economics, 8*(3), 249–258.

Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography, 28*(1), 31–56.

Beaudry, C., & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy, 38*(2), 318–337.

Berliant, M., & Fujita, M. (2011). The dynamics of knowledge diversity and economic growth. *Southern Economic Journal, 77*(4), 856–884.

Bettencourt, L. M., Lobo, J., & Strumsky, D. (2007). Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size. *Research Policy, 36*, 107–120.

Boschma, R. A., & Fritsch, M. (2009). Creative class and regional growth: Empirical evidence from seven European countries. *Economic Geography, 85*(4), 391–423.

Boschma, R., & Iammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. *Economic Geography, 85*(3), 289–311.

Boschma, R., Minondo, A., & Navarro, M. (2012). Related variety and regional growth in Spain. *Papers in Regional Science, 91*(2), 241–256.

Breschi, S. (1999). Spatial patterns of innovation. In A. Gambardella & F. Malerba (Eds.), *The organisation of economic innovation in Europe* (pp. 71–102). Cambridge: Cambridge University Press.

Breschi, S., Lissoni, F., & Malerba, F. (2003). Knowledge-relatedness in firm technological diversification. *Research Policy, 32*(1), 69–87.

Cameron, C., & Trivedi, P. (2008). *Applied Microeconometrics using STATA*. New York: STATA Press.

Castellani, D. (2002). Export behaviour and productivity growth: Evidence from Italian manufacturing firms. *Review of World Economics, 138*(4), 605–628.

Clerides, S., Lach, S., & Tybout, J. (1998). Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco. *The Quarterly Journal of Economics, 113*(3), 903–947.

Crescenzi, R., Rodriguez-Pose, A., & Storper, M. (2012). The territorial dynamics of innovation in China and India. *Journal of Economic Geography, 12*, 1055–1085.

De Groot, H. L., Poot, J., & Smit, M. J. (2009). Agglomeration externalities, innovation and regional growth: Theoretical perspectives and meta-analysis. In R. Capello & P. Nijkamp (Eds.), *Handbook of regional growth and development theories* (pp. 256–281). Northampton MA: Edward Elgar Publishing.

Duranton, G., & Puga, D. (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *The American Economic Review, 9*(5), 1454–1477.

Ejermo, O. (2005). Technological diversity and Jacobs' externality hypothesis revisited. *Growth and Change, 36*(2), 167–195.

Essletzbichler, J. (2007). Diversity, stability and regional growth in the United States 1975–2002. In K. Frenken (Ed.), *Applied evolutionary economics and economic geography* (pp. 203–229). Cheltenham: Edward Edgar Publishing.

Feldman, M. P. (1994). *The geography of innovation*. Dordrecht: Kluwer Academic.

Feldman, M. P. (2003). Location and innovation: The new economic geography of innovation, spillovers, and agglomeration. In G. L. Clark & M. P. Feldman (Eds.), *The Oxford handbook of economic geography* (pp. 373–394). Oxford: Oxford University Press.

Feldman, M. P., & Audretsch, D. B. (1999). Innovation in cities: Science-based diversity, specialization and localized competition. *European Economic Review, 43*, 409–429.

Feldman, M. P., & Florida, R. (1994). The geographic sources of innovation: Technological infrastructure and product innovation in the United States. *Annals of the Association of American Geographers, 84*, 210–229.

Feldman, M. P., & Kogler, D. F. (2010). Stylized facts in the geography of innovation. In B. Hall & N. Rosenberg (Eds.), *Handbook of economics of technical change* (pp. 381–410). Oxford: Elsevier.

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies, 41*(5), 685–697.

Fritsch, M., & Slavtchev, V. (2007). Universities and innovation in space. *Industry and Innovation, 14*(2), 201–218.

Glaeser, E., Kallal, H., Scheinkam, J., & Shleifer, A. (1992). Growth in cities. *The Journal of Political Economy, 100*, 1126–1152.

Gordon, I. R., & Mccann, P. (2000). Industrial clusters, complexes, agglomeration and/or social networks? *Urban Studies, 37*, 513–532.

Grant, R. M. (1996). Prospering in dynamically-competitive environments: Organizational capability as knowledge integration. *Organization Science, 7*(4), 375–387.

Gråsjö U. (2006) Spatial spillovers of knowledge production—An accessibility approach. JIBS Dissertation series No. 034, Jönköping.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *The Bell Journal of Economics, 10*(1), 92–116.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature, 28*(4), 1661–1707.

Hipp, C., & Grupp, H. (2005). Innovation in the service sector: The demand for service-specific innovation measurement concepts and typologies. *Research Policy, 34*, 517–535.

Jacobs, J. (1969). *The economy of cities*. New York: Vintage.

Jaffe, A. B. (1989). Real effects of academic research. *American Economic Review, 79*, 957–970.

Jaffe, A. B., & Trajtenberg, M. (2005). *Patents, citations, and innovations: A window on the knowledge economy*. Massachusetts: MIT Press Books.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics, 108*(3), 577–598.

Karlsson, C., & Johansson, B. (2006). Dynamic and entrepreneurship in a knowledge-based economy. In C. Karlsson, B. Johansson, & R. Stough (Eds.), *Entrepreneurship and dynamics in the knowledge economy* (pp. 12–46). New York: Routledge.

Karlsson, C., & Nyström, K. (2011). Knowledge accessibility and new firm formation. In S. Desai, P. Nijkamp, & R. Stough (Eds.), *New directions in regional economic development* (pp. 174–198). Cheltenham: Edward Elgar Publishing.

Keller, W. (2004). International technology diffusion. *Journal of Economic Literature, 42*(3), 752–782.

Krugman, P. R. (1991). *Geography and trade*. Cambridge: MIT Press.

Lee, S. Y., Florida, R., & Gates, G. (2010). Innovation, human capital, and diversity. *International Review of Public Administration, 14*(3), 13–24.

Lucas, R. (1988). On the mechanics of economic growth. *Journal of Monetary Economics, 22*, 3–42.

Malerba, F., & Orsenigo, L. (1997). Technological regimes and sectoral patterns of innovative activities. *Industrial and Corporate Change, 6*(1), 83–117.

Maskell, P. (2001). Towards a knowledge-based theory of the geographical cluster. *Industrial and Corporate Change, 10*, 921–943.

Maskell, P., & Malmberg, A. (1999). Localised learning and industrial competitiveness. *Cambridge Journal of Economics, 23*(2), 167–186.

Maskell, P., & Malmberg, A. (2002). The elusive concept of localization economies: Towards a knowledge-based theory of spatial clustering. *Environment and Planning A, 34*, 429–449.

Mesquita, L. F. (2007). Starting over when the bickering never ends: Rebuilding aggregate trust among clustered firms through trust facilitators. *Academy of Management Review, 32*, 72–91.

Neffke, F., Henning, M., & Boschma, R. (2011). How do regions diversify over time? Industry relatedness and the development industry relatedness and the development. *Economic Geography, 87*(3), 237–265.

NUTEK. (2005). *Årsbok*. Stockholm: The Swedish Agency for Economic and Regional Growth.

Nooteboom, B. (2000). *Learning and innovation in organizations and economies*. Oxford: Oxford University Press.

Paci, R., & Usai, S. (1999). Externalities, knowledge spillovers and the spatial distribution of innovation. *GeoJournal, 49*, 381–390.

Ponds, R., Van Oort, F., & Frenken, K. (2010). Innovation, spillovers and university–industry collaboration: An extended knowledge production function approach. *Journal of Economic Geography, 10*, 231–255.

Rodan, S., & Galunic, C. (2004). More than network structure: How knowledge heterogeneity influences managerial performance and innovativeness. *Strategic Management Journal, 25*, 541–562.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy, 94*, 1002–1037.

Saviotti, P. P. (1996). *Technological evolution, variety and the economy*. Cheltenham: Edward Elgar Publishing.

Saxenian, A. (1996). Regional advantage: Culture and competition in silicon valley and route 128. Cambridge: Harvard University Press.

Scherer, F. M. (1983). The propensity to patent. *International Journal of Industrial Organization, 1*, 107–128.

Schumpeter, J. (1934). *The theory of economic development*. Cambridge, MA: Harvard University Press.

Tallman, S., Jenkins, M., Henry, N., & Pinch, S. (2004). Knowledge, clusters and competitive advantage. *Academy of Management Review, 29*, 258–271.

Utterback, J. M., & Abernathy, W. J. (1975). A dynamic model of process and product innovation. *Omega, 3*(6), 639–656.

Vernon, R. (1966). International investment and international trade in the product cycle. *The Quarterly Journal of Economics, 2*, 190–207.